

SEQUENCE LISTING

SEQ ID NO:1

EBOGP1/MBGGP2 nucleotide sequence

```

      10      20      30      40      50
atgggcgcta caggaatatt gcagttacct cgtgatcgat tcaagaggac

      60      70      80      90     100
atcattcttt ctttgggtaa ttatcctttt ccaaagaaca ttttccatcc

     110     120     130     140     150
cacttggagt catccacaat agcacattac aggttagtga tgtcgacaaa

     160     170     180     190     200
ctagtttgtc gtgacaaaact gtcattccaca aatcaattga gatcagttgg

     210     220     230     240     250
actgaatctc gaagggaatg gagtggcaac tgacgtgccca tctgcaacta

     260     270     280     290     300
aaagatgggg cttcaggtcc ggtgtcccac caaaggtggt caattatgaa

     310     320     330     340     350
gctggtgaat gggctgaaaa ctgctacaat cttgaaatca aaaaacctga

     360     370     380     390     400
cgggagtgag tgtctaccag cagcgccaga cgggattcgg ggcttcccc

     410     420     430     440     450
ggtgccgcta tgtgcacaaa gtatcaggaa cgggaccgtg tgccggagac

     460     470     480     490     500
tttgcttcc ataaagaggg tgctttcttc ctgtatgatc gacttgcttc

     510     520     530     540     550
cacagttatc taccgaggaa cgactttcgc tgaaggtgtc gttgcatttc

     560     570     580     590     600
tgatactgcc ccaagctaag aaggacttct tcagctcaca ccccttgaga

     610     620     630     640     650
gagccggtca atgcaacgga ggaccctctt agtggctact attctaccac

     660     670     680     690     700
aattagatat caggctaccg gttttggaac caatgagaca gagtacttgt

     710     720     730     740     750
tcgaggttga caatttgacc tacgtccaac ttgaatcaag attcacacca

     760     770     780     790     800
cagtttctgc tccagctgaa tgagacaata tatacaagtg ggaaaaggag

```

EBOGP1/MBGGP2

```

      810      820      830      840      850
caataccacg ggaaaaactaa tttggaaggt caaccccgaa attgatacaa

      860      870      880      890      900
caatcgggga gtgggccttc tgggaaacta aaaaaaacct cactagaaaa

      910      920      930      940      950
attcgcagtg aagagttgtc tttcacagtt gtatcaaacg gagccaaaaa

      960      970      980      990     1000
catcagtggc cagagtccgg cgcgaaactc ttccgaccca gggaccaaca

     1010     1020     1030     1040     1050
caacaactga agaccacaaa atcatggctt cagaaaattc ctctgcaatg

     1060     1070     1080     1090     1100
gttcaagtgc acagtcaagg aagggaagct gcagtgtcgc atctaacaac

     1110     1120     1130     1140     1150
ccttgccaca atctccacga gtccccaatc cctcacaacc aaaccaggtc

     1160     1170     1180     1190     1200
cggacaacag cacccataat acaccgtgt ataaacttga catctctgag

     1210     1220     1230     1240     1250
gcaactcaag ttgaacaaca tcaccgcaga acagacaacg acagcacagc

     1260     1270     1280     1290     1300
ctccgacact ccctctgcc aagaccgcgc cggaccccca aaagcagaga

     1310     1320     1330     1340     1350
acaccaacac gagcaagagc actgacttcc tggaccccg caccacaaca

     1360     1370     1380     1390     1400
agtcccaaaa accacagcga gaccgctggc aacaacaaca ctcatcacca

     1410     1420     1430     1440     1450
agataccgga gaagagagtg ccagcagcgg gaagctaggc ttaattacca

     1460     1470     1480     1490     1500
atactattgc tggagtgcga ggactgatca caggcgggag aagaactcga

     1510     1520     1530     1540     1550
cgatcgatcc tctggaggga aggcgacatg ttcccttttc tggatgggtt

     1560     1570     1580     1590     1600
aataaatgct ccaattgatt ttgaccagat tccaaataca aaaacaatct

     1610     1620     1630     1640     1650
ttgatgaatc ctctagttct ggtgcctcgg ctgaggaaga tcaacatgcc

     1660     1670     1680     1690     1700
tcccccaata ttagtttaac tttatcttat tttcctaata taaatgagaa

```

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622
 623
 624
 625
 626
 627
 628
 629
 630
 631
 632
 633
 634
 635
 636
 637
 638
 639
 640
 641
 642
 643
 644
 645
 646
 647
 648
 649
 650
 651
 652
 653
 654
 655
 656
 657
 658
 659
 660
 661
 662
 663
 664
 665
 666
 667
 668
 669
 670
 671
 672
 673
 674
 675
 676
 677
 678
 679
 680
 681
 682
 683
 684
 685
 686
 687
 688
 689
 690
 691
 692
 693
 694
 695
 696
 697
 698
 699
 700
 701
 702
 703
 704
 705
 706
 707
 708
 709
 710
 711
 712
 713
 714
 715
 716
 717
 718
 719
 720
 721
 722
 723
 724
 725
 726
 727
 728
 729
 730
 731
 732
 733
 734
 735
 736
 737
 738
 739
 740
 741
 742
 743
 744
 745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755
 756
 757
 758
 759
 760
 761
 762
 763
 764
 765
 766
 767
 768
 769
 770
 771
 772
 773
 774
 775
 776
 777
 778
 779
 780
 781
 782
 783
 784
 785
 786
 787
 788
 789
 790
 791
 792
 793
 794
 795
 796
 797
 798
 799
 800
 801
 802
 803
 804
 805
 806
 807
 808
 809
 810
 811
 812
 813
 814
 815
 816
 817
 818
 819
 820
 821
 822
 823
 824
 825
 826
 827
 828
 829
 830
 831
 832
 833
 834
 835
 836
 837
 838
 839
 840
 841
 842
 843
 844
 845
 846
 847
 848
 849
 850
 851
 852
 853
 854
 855
 856
 857
 858
 859
 860
 861
 862
 863
 864
 865
 866
 867
 868
 869
 870
 871
 872
 873
 874
 875
 876
 877
 878
 879
 880
 881
 882
 883
 884
 885
 886
 887
 888
 889
 890
 891
 892
 893
 894
 895
 896
 897
 898
 899
 900
 901
 902
 903
 904
 905
 906
 907
 908
 909
 910
 911
 912
 913
 914
 915
 916
 917
 918
 919
 920
 921
 922
 923
 924
 925
 926
 927
 928
 929
 930
 931
 932
 933
 934
 935
 936
 937
 938
 939
 940
 941
 942
 943
 944
 945
 946
 947
 948
 949
 950
 951
 952
 953
 954
 955
 956
 957
 958
 959
 960
 961
 962
 963
 964
 965
 966
 967
 968
 969
 970
 971
 972
 973
 974
 975
 976
 977
 978
 979
 980
 981
 982
 983
 984
 985
 986
 987
 988
 989
 990
 991
 992
 993
 994
 995
 996
 997
 998
 999
 1000

```

      1710      1720      1730      1740      1750
cactgcctac tctggagaaa atgagaatga ttgtgatgca gagttaagaa

      1760      1770      1780      1790      1800
tttggagcgt tcaggaggat gacctggccg cagggctcag ttggataccg

      1810      1820      1830      1840      1850
ttttttggcc ctggaattga aggactttac actgctgttt taattaaanaa

      1860      1870      1880      1890      1900
tcaaaaacaat ttggtctgca ggttgaggcg tctagccaat caaactgcca

      1910      1920      1930      1940      1950
aatccttgga actcttattg agagtcacaa ctgaggaaag aacattctcc

      1960      1970      1980      1990      2000
ttaatcaata gacatgctat tgactttcta ctcacaagat ggggaggaac

      2010      2020      2030      2040      2050
atgcaaagtg cttggacctg attgttgcac cgggatagaa gacttgtcca

      2060      2070      2080      2090      2100
aaaatatttc agagcaaatt gaccaaatta aaaaggacga acaaaaagag

      2110      2120      2130      2140      2150
gggactgggt ggggtctggg tggtaaattg tggacatccg actgggggtg

      2160      2170      2180      2190      2200
tcttactaac ttgggcattt tgctactatt atccatagct gtcttgattg

      2210      2220      2230      2240      2250
ctctatcctg tatttgtcgt atctttacta aatatatcgg ataacggaat

tc

```

SEQ ID NO:2

EBOGP1/MBGGP2 amino acid sequence

MetGlyValThrGlyIleLeuGlnLeuProArgAspArgPheLys

ArgThrSerPhePheLeuTrpValIleIleLeuPheGlnArgThrPhe

SerIleProLeuGlyValIleHisAsnSerThrLeuGlnValSerAsp

ValAspLysLeuValCysArgAspLysLeuSerSerThrAsnGlnLeuArg

SerValGlyLeuAsnLeuGluGlyAsnGlyValAlaThrAspValProSer

AlaThrLysArgTrpGlyPheArgSerGlyValProProLysValVal

AsnTyrGluAlaGlyGluTrpAlaGluAsnCysTyrAsnLeuGluIleLys

LysProAspGlySerGluCysLeuProAlaAlaProAspGlyIleArgGly

PheProArgCysArgTyrValHisLysValSerGlyThrGlyProCys
AlaGlyAspPheAlaPheHisLysGluGlyAlaPhePheLeuTyrAspArg
LeuAlaSerThrValIleTyrArgGlyThrThrPheAlaGluGlyValVal
AlaPheLeuIleLeuProGlnAlaLysLysAspPhePheSerSerHis
ProLeuArgGluProValAsnAlaThrGluAspProSerSerGlyTyrTyr
SerThrThrIleArgTyrGlnAlaThrGlyPheGlyThrAsnGluThrGlu
TyrLeuPheGluValAspAsnLeuThrTyrValGlnLeuGluSerArg
PheThrProGlnPheLeuLeuGlnLeuAsnGluThrIleTyrThrSerGly
LysArgSerAsnThrThrGlyLysLeuIleTrpLysValAsnProGluIle
AspThrThrIleGlyGluTrpAlaPheTrpGluThrLysLysAsnLeu
ThrArgLysIleArgSerGluGluLeuSerPheThrValValSerAsnGly
AlaLysAsnIleSerGlyGlnSerProAlaArgThrSerSerAspProGly
ThrAsnThrThrThrGluAspHisLysIleMetAlaSerGluAsnSer
SerAlaMetValGlnValHisSerGlnGlyArgGluAlaAlaValSerHis
LeuThrThrLeuAlaThrIleSerThrSerProGlnSerLeuThrThrLys
ProGlyProAspAsnSerThrHisAsnThrProValTyrLysLeuAsp
IleSerGluAlaThrGlnValGluGlnHisHisArgArgThrAspAsnAsp
SerThrAlaSerAspThrProSerAlaThrThrAlaAlaGlyProProLys
AlaGluAsnThrAsnThrSerLysSerThrAspPheLeuAspProAla
ThrThrThrSerProGlnAsnHisSerGluThrAlaGlyAsnAsnAsnThr
HisHisGlnAspThrGlyGluGluSerAlaSerSerGlyLysLeuGlyLeu
IleThrAsnThrIleAlaGlyValAlaGlyLeuIleThrGlyGlyArg
ArgThrArgArgSerIleLeuTrpArgGluGlyAspMetPheProPheLeu
AspGlyLeuIleAsnAlaProIleAspPheAspProValProAsnThrLys
ThrIlePheAspGluSerSerSerSerGlyAlaSerAlaGluGluAsp
GlnHisAlaSerProAsnIleSerLeuThrLeuSerTyrPheProAsnIle
AsnGluAsnThrAlaTyrSerGlyGluAsnGluAsnAspCysAspAlaGlu

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

LeuArgIleTrpSerValGlnGluAspAspLeuAlaAlaGlyLeuSer
 TrpIleProPhePheGlyProGlyIleGluGlyLeuTyrThrAlaValLeu
 IleLysAsnGlnAsnAsnLeuValCysArgLeuArgArgLeuAlaAsnGln
 ThrAlaLysSerLeuGluLeuLeuLeuArgValThrThrGluGluArg
 ThrPheSerLeuIleAsnArgHisAlaIleAspPheLeuLeuThrArgTrp
 GlyGlyThrCysLysValLeuGlyProAspCysCysIleGlyIleGluAsp
 LeuSerLysAsnIleSerGluGlnIleAspGlnIleLysLysAspGlu
 GlnLysGluGlyThrGlyTrpGlyLeuGlyGlyLysTrpTrpThrSerAsp
 TrpGlyValLeuThrAsnLeuGlyIleLeuLeuLeuLeuSerIleAlaVal
 LeuIleAlaLeuSerCysIleCysArgIlePheThrLysTyrIleGly*

SEQ ID NO:3

MBGGP1/EBGP2 nucleotide sequence

10	20	30	40	50
atgaagacca	catgtttcct	tatcagtctt	atcttaattc	aagggacaaa
60	70	80	90	100
aaatctcccc	atttttagaga	tagctagtaa	taatcaaccc	caaaatgtgg
110	120	130	140	150
attcggatg	ctccggaact	ctccagaaga	cagaagacgt	ccatctgatg
160	170	180	190	200
ggattcacac	tgagtgggca	aaaagttgct	gattcccctt	tggaggcatc
210	220	230	240	250
caagcgatgg	gcttttcagga	caggtgtacc	tccaagaat	gttgagtaca
260	270	280	290	300
cagaggggga	ggaagccaaa	acatgctaca	atataagtgt	aacggatccc
310	320	330	340	350
tctggaaaat	ccttgctgtt	agatcctcct	accaacatcc	gtgactatcc
360	370	380	390	400
gaaatgcaaa	actatccatc	atattcaagg	tcaaaaccct	catgcacagg
410	420	430	440	450
ggatcgccct	tcattttatgg	ggagcatttt	ttctgtatga	tcgcattgcc
460	470	480	490	500
tccacaacaa	tgtaccgagg	caaagtcttc	actgaaggga	acatagcagc

MBGGP1/EBGP2 nucleotide sequence

```

510      520      530      540      550
tatgattgtc aataagacag tgcacaaaat gattttctcg cggcaaggac

560      570      580      590      600
aagggtaccg tcatatgaat ctgacttcta ctaataaata ttggacaagt

610      620      630      640      650
agtaacggaa cgcaaacgaa tgacactgga tgtttcggcg ctcttcaaga

660      670      680      690      700
atacaattct acaaagaacc aaacatgtgc tccgtccaaa atacctccac

710      720      730      740      750
cactgcccac agcccgtccg gagatcaaac tcacaagcac cccaactgat

760      770      780      790      800
gccaccaaac tcaataccac ggacccaagc agtgatgatg aggacctcgc

810      820      830      840      850
aacatccggc tcaggggtccg gagaacgaga accccacaca acttctgatg

860      870      880      890      900
cggtcaccaa gcaagggctt tcatcaacaa tgccaccac tccctcacca

910      920      930      940      950
caaccaagca cgccacagca aggaggaaac aacacaaacc attcccaaga

960      970      980      990      1000
tgctgtgact gaactagaca aaaataacac aactgcacaa ccgtccatgc

1010     1020     1030     1040     1050
ccctcataa cactaccaca atctctacta acaacacctc caaacacaa

1060     1070     1080     1090     1100
ttcagcactc tctctgcacc attacaaaac accaccaatg acaacacaca

1110     1120     1130     1140     1150
gagcacaatc actgaaaatg agcaaaccag tgccccctcg ataacaacc

1160     1170     1180     1190     1200
tgctccaac gggaaatccc accacagcaa agagcaccag cagcaaaaaa

1210     1220     1230     1240     1250
ggccccgcca caacggcacc aaacacgaca aatgagcatt tcaccagtcc

1260     1270     1280     1290     1300
tccccccacc cccagctcga ctgcacaaca tcttgtatat ttcagaagaa

1310     1320     1330     1340     1350
agcgatcggc aattgtcaat gctcaacca aatgcaacc taatttacat

1360     1370     1380     1390     1400
tactggacta ctcaggatga aggtgctgca atcggactgg cctggatacc

```

```

      1410      1420      1430      1440      1450
atattttcggg ccagcagccg agggaattta catagagggg ctaatgcaca

      1460      1470      1480      1490      1500
atcaagatgg tttaatctgt gggttgagac agctggccaa cgagacgact

      1510      1520      1530      1540      1550
caagctcttc aactgttcct gagagccaca actgagctac gcaccttttc

      1560      1570      1580      1590      1600
aatcctcaac cgtaaggcaa ttgatttctt gctgcagcga tggggcggca

      1610      1620      1630      1640      1650
catgccacat tctgggaccg gactgctgta tcgaaccaca tgattggacc

      1660      1670      1680      1690      1700
aagaacataa cagacaaaat tgatcagatt attcatgatt ttgttgataa

      1710      1720      1730      1740      1750
aacccttccg gaccagggggg acaatgacaa ttgggtggaca ggatggagac

      1760      1770      1780      1790      1800
aatggatacc ggcaggtatt ggagttacag gcgttataat tgcagttatc

      1810      1820      1830      1840
gctttattct gtatatgcaa atttgtcttt tagttgaatt c

```

SEQ ID NO:4

MBG1/EBO2 amino acid sequence

MetLysThrThrCysPheLeuIleSerLeuIleLeuIleGlnGlyThr
 LysAsnLeuProIleLeuGluIleAlaSerAsnAsnGlnProGln
 AsnValAspSerValCysSerGlyThrLeuGlnLysThrGluAspVal
 HisLeuMetGlyPheThrLeuSerGlyGlnLysValAlaAspSerProLeu
 GluAlaSerLysArgTrpAlaPheArgThrGlyValProProLysAsn
 ValGluTyrThrGluGlyGluGluAlaLysThrCysTyrAsnIleSerVal
 ThrAspProSerGlyLysSerLeuLeuLeuAspProProThrAsnIleArg
 AspTyrProLysCysLysThrIleHisHisIleGlnGlyGlnAsnPro
 HisAlaGlnGlyIleAlaLeuHisLeuTrpGlyAlaPhePheLeuTyrAsp
 ArgIleAlaSerThrThrMetTyrArgGlyLysValPheThrGluGlyAsn
 IleAlaAlaMetIleValAsnLysThrValHisLysMetIlePheSer

ArgGlnGlyGlnGlyTyrArgHisMetAsnLeuThrSerThrAsnLysTyr
 TrpThrSerSerAsnGlyThrGlnThrAsnAspThrGlyCysPheGlyAla
 LeuGlnGluTyrAsnSerThrLysAsnGlnThrCysAlaProSerLys
 IleProProProLeuProThrAlaArgProGluIleLysLeuThrSerThr
 ProThrAspAlaThrLysLeuAsnThrThrAspProSerSerAspAspGlu
 AspLeuAlaThrSerGlySerGlySerGlyGluArgGluProHisThr
 ThrSerAspAlaValThrLysGlnGlyLeuSerSerThrMetProProThr
 ProSerProGlnProSerThrProGlnGlnGlyGlyAsnAsnThrAsnHis
 SerGlnAspAlaValThrGluLeuAspLysAsnAsnThrThrAlaGln
 ProSerMetProProHisAsnThrThrThrIleSerThrAsnAsnThrSer
 LysHisAsnPheSerThrLeuSerAlaProLeuGlnAsnThrThrAsnAsp
 AsnThrGlnSerThrIleThrGluAsnGluGlnThrSerAlaProSer
 IleThrThrLeuProProThrGlyAsnProThrThrAlaLysSerThrSer
 SerLysLysGlyProAlaThrThrAlaProAsnThrThrAsnGluHisPhe
 ThrSerProProProThrProSerSerThrAlaGlnHisLeuValTyr
 PheArgArgLysArgSerAlaIleValAsnAlaGlnProLysCysAsnPro
 AsnLeuHisTyrTrpThrThrGlnAspGluGlyAlaAlaIleGlyLeuAla
 TrpIleProTyrPheGlyProAlaAlaGluGlyIleTyrIleGluGly
 LeuMetHisAsnGlnAspGlyLeuIleCysGlyLeuArgGlnLeuAlaAsn
 GluThrThrGlnAlaLeuGlnLeuPheLeuArgAlaThrThrGluLeuArg
 ThrPheSerIleLeuAsnArgLysAlaIleAspPheLeuLeuGlnArg
 TrpGlyGlyThrCysHisIleLeuGlyProAspCysCysIleGluProHis
 AspTrpThrLysAsnIleThrAspLysIleAspGlnIleIleHisAspPhe
 ValAspLysThrLeuProAspGlnGlyAspAsnAspAsnTrpTrpThr
 GlyTrpArgGlnTrpIleProAlaGlyIleGlyValThrGlyValIleIle
 AlaValIleAlaLeuPheCysIleCysLysPheValPhe*

2017-11-11 11:11:11

SEQ ID NO:5

MUSGP1/RVNGP2 nucleotide sequence

Sequence Range: 1 to 2046

```

      10      20      30      40      50
atgaagacca catgtttcct tatcagtctt atcttaattc aagggacaaa

      60      70      80      90     100
aaatctcccc attttagaga tagctagtaa taatcaaccc caaaatgtgg

     110     120     130     140     150
attcggtatg ctccggaact ctccagaaga cagaagacgt ccatctgatg

     160     170     180     190     200
ggattcacac tgagtgggca aaaagttgct gattccccct tggaggcatc

     210     220     230     240     250
caagcgatgg gctttcagga caggtgtacc tcccaagaat gttgagtaca

     260     270     280     290     300
cagaggggga ggaagccaaa acatgctaca atataagtgt aacggatccc

     310     320     330     340     350
tctggaaaat ccttgctgtt agatcctcct accaacaatcc gtgactatcc

     360     370     380     390     400
gaaatgcaaa actatccatc atattcaagg tcaaaaccct catgcacagg

     410     420     430     440     450
ggatcgccct tcatttatgg ggagcatttt ttctgtatga tcgcattgcc

     460     470     480     490     500
tccacaacaa tgtaccgagg caaagtcttc actgaaggga acatagcagc

     510     520     530     540     550
tatgattgtc aataagacag tgcacaaaat gatcttctcg cggcaaggac

     560     570     580     590     600
aagggtaccg tcatatgaat ctgacttcta ctaataaata ttggacaagt

     610     620     630     640     650
agtaacggaa cgcaaacgaa tgacactgga tgtttcggcg ctcttcaaga

     660     670     680     690     700
atacaattct acaaagaacc aaacatgtgc tccgtccaaa atacctccac

     710     720     730     740     750
cactgcccac agcccgtccg gagatcaaac tcacaagcac cccaactgat

     760     770     780     790     800
gccaccaaac tcaataccac ggaccaagc agtgatgatg aggacctcgc

     810     820     830     840     850

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046

aacatccggc	tcaggggtccg	gagaacgaga	acccacaca	acttctgatg
860	870	880	890	900
cggtcaccaa	gcaaggggctt	tcatcaacaa	tgccaccac	tccctcacca
910	920	930	940	950
caaccaagca	cgccacagca	aggaggaaaac	aacacaaacc	attcccaaga
960	970	980	990	1000
tgctgtgact	gaactagaca	aaaataaacac	aactgcacaa	ccgtccatgc
1010	1020	1030	1040	1050
cccctcataa	cactaccaca	atctctacta	acaacacctc	caaacacaac
1060	1070	1080	1090	1100
ttcagcactc	tctctgcacc	attacaaaac	accaccaatg	acaacacaca
1110	1120	1130	1140	1150
gagcacaatc	actgaaaatg	agcaaaccag	tgccccctcg	ataacaaccc
1160	1170	1180	1190	1200
tgcttccaac	gggaaatccc	accacagcaa	agagcaccag	cagcaaaaaa
1210	1220	1230	1240	1250
ggccccgcca	caacggcacc	aaacacgaca	aatgagcatt	tcaccagtcc
1260	1270	1280	1290	1300
ttccccccacc	cccagctcga	ctgcacaaca	tcttgtatat	ttcagaagaa
1310	1320	1330	1340	1350
agcgatcgat	tttctggaaa	gaaggtgata	tattcccgtt	tttagatggg
1360	1370	1380	1390	1400
ttaataaata	ctgaaattga	ttttgatcca	atcccaaaca	cagaaacaat
1410	1420	1430	1440	1450
ctttgatgaa	tctcccgagt	ttaatacttc	aactaatgag	gaacaacaca
1460	1470	1480	1490	1500
ctcccccgaa	tatcagttta	actttctctt	attttctctga	taaaaatgga
1510	1520	1530	1540	1550
gatactgcct	actctgggga	aaacgagaat	gattgtgatg	cagagttgag
1560	1570	1580	1590	1600
gatttggagt	gtgcaggagg	acgatttggc	ggcaggggctt	agctggatac
1610	1620	1630	1640	1650
catttttttg	ccctggaatc	gaaggactct	atactgccgg	tttaatcaaa
1660	1670	1680	1690	1700
aatcagaaca	atttagtttg	taggttgagg	cgcttagcta	atcaaactgc
1710	1720	1730	1740	1750

```

taaatccttg gagctcttgt taagggtcac aaccgaggaa aggacatttt
      1760      1770      1780      1790      1800
ccttaatcaa taggcatgca attgactttt tgcttacgag gtggggcgga
      1810      1820      1830      1840      1850
acatgcaagg tgctaggacc tgattgttgc ataggaatag aagatctatc
      1860      1870      1880      1890      1900
taaaaatatc tcagaacaaa tcgacaaaat cagaaaggat gaacaaaagg
      1910      1920      1930      1940      1950
aggaaaactgg ctgggggtcta ggtggcaaat ggtggacatc tgactggggt
      1960      1970      1980      1990      2000
gttctcacca atttgggcat cctgctacta ttatctatag ctgttctgat
      2010      2020      2030      2040
tgctctgtcc tgtatctgtc gtatcttcac taaatacatt ggatga

```

SEQ ID NO:6

MUSGP1/RVNGP2 amino acid sequence

MetLysThrThrCysPheLeuIleSerLeuIleLeuIleGlnGlyThr
 LysAsnLeuProIleLeuGluIleAlaSerAsnAsnGlnProGln
 AsnValAspSerValCysSerGlyThrLeuGlnLysThrGluAspVal
 HisLeuMetGlyPheThrLeuSerGlyGlnLysValAlaAspSerProLeu
 GluAlaSerLysArgTrpAlaPheArgThrGlyValProProLysAsn
 ValGluTyrThrGluGlyGluGluAlaLysThrCysTyrAsnIleSerVal
 ThrAspProSerGlyLysSerLeuLeuLeuAspProProThrAsnIleArg
 AspTyrProLysCysLysThrIleHisHisIleGlnGlyGlnAsnPro
 HisAlaGlnGlyIleAlaLeuHisLeuTrpGlyAlaPhePheLeuTyrAsp
 ArgIleAlaSerThrThrMetTyrArgGlyLysValPheThrGluGlyAsn
 IleAlaAlaMetIleValAsnLysThrValHisLysMetIlePheSer
 ArgGlnGlyGlnGlyTyrArgHisMetAsnLeuThrSerThrAsnLysTyr
 TrpThrSerSerAsnGlyThrGlnThrAsnAspThrGlyCysPheGlyAla
 LeuGlnGluTyrAsnSerThrLysAsnGlnThrCysAlaProSerLys
 IleProProProLeuProThrAlaArgProGluIleLysLeuThrSerThr

ProThrAspAlaThrLysLeuAsnThrThrAspProSerSerAspAspGlu
AspLeuAlaThrSerGlySerGlySerGlyGluArgGluProHisThr
ThrSerAspAlaValThrLysGlnGlyLeuSerSerThrMetProProThr
ProSerProGlnProSerThrProGlnGlnGlyGlyAsnAsnThrAsnHis
SerGlnAspAlaValThrGluLeuAspLysAsnAsnThrThrAlaGln
ProSerMetProProHisAsnThrThrThrIleSerThrAsnAsnThrSer
LysHisAsnPheSerThrLeuSerAlaProLeuGlnAsnThrThrAsnAsp
AsnThrGlnSerThrIleThrGluAsnGluGlnThrSerAlaProSer
IleThrThrLeuProProThrGlyAsnProThrThrAlaLysSerThrSer
SerLysLysGlyProAlaThrThrAlaProAsnThrThrAsnGluHisPhe
ThrSerProProProThrProSerSerThrAlaGlnHisLeuValTyr
PheArgArgLysArgSerIlePheTrpLysGluGlyAspIlePheProPhe
LeuAspGlyLeuIleAsnThrGluIleAspPheAspProIleProAsnThr
GluThrIlePheAspGluSerProSerPheAsnThrSerThrAsnGlu
GluGlnHisThrProProAsnIleSerLeuThrPheSerTyrPheProAsp
LysAsnGlyAspThrAlaTyrSerGlyGluAsnGluAsnAspCysAspAla
GluLeuArgIleTrpSerValGlnGluAspAspLeuAlaAlaGlyLeu
SerTrpIleProPhePheGlyProGlyIleGluGlyLeuTyrThrAlaGly
LeuIleLysAsnGlnAsnAsnLeuValCysArgLeuArgArgLeuAlaAsn
GlnThrAlaLysSerLeuGluLeuLeuLeuArgValThrThrGluGlu
ArgThrPheSerLeuIleAsnArgHisAlaIleAspPheLeuLeuThrArg
TrpGlyGlyThrCysLysValLeuGlyProAspCysCysIleGlyIleGlu
AspLeuSerLysAsnIleSerGluGlnIleAspLysIleArgLysAsp
GluGlnLysGluGluThrGlyTrpGlyLeuGlyGlyLysTrpTrpThrSer
AspTrpGlyValLeuThrAsnLeuGlyIleLeuLeuLeuLeuSerIleAla
ValLeuIleAlaLeuSerCysIleCysArgIlePheThrLysTyrIleGly*

SEQ ID NO:7

RVNGP1/MUSGP2 nucleotide sequence

```

      10      20      30      40      50
atgaagacca tatattttct gattagtctc attttaatcc aaagtataaa

      60      70      80      90     100
aactctccct gtttttagaaa ttgctagtaa cagccaacct caagatgtag

     110     120     130     140     150
attcagtgtg ctccggaacc ctccaaaaga cagaagatgt tcatctgatg

     160     170     180     190     200
ggattttacac tgagtgggca aaaagttgct gattccccctt tggaagcatc

     210     220     230     240     250
taaacgatgg gctttcagga cagggtgtcc tccaagaac gttgagtata

     260     270     280     290     300
cggaaggaga agaagccaaa acatgttaca atataagtggt aacagaccct

     310     320     330     340     350
tctggaaaat ccttgctgct ggatcctccc agtaatatcc gcgattaccc

     360     370     380     390     400
taaatgtaaa actgttcatc atattcaagg tcaaaaccct catgcacagg

     410     420     430     440     450
ggattgccct ccatttggtg ggggcatttt tcttgatga tcgcgttgcc

     460     470     480     490     500
tctacaacaa tgtaccgagg caaggtcttc actgaaggaa atatagcagc

     510     520     530     540     550
tatgattggt aataagacag ttcacagaat gattttttct aggcaaggac

     560     570     580     590     600
aaggttatcg tcacatgaac ttgacctcca ccaataaata ttggacaagc

     610     620     630     640     650
agcaatgaaa cgcagagaaa tgatacggga tgttttggca tcctccaaga

     660     670     680     690     700
atacaactcc acaaacaatc aaacatgccc tccatctctt aaacctccat

     710     720     730     740     750
ccctgcccac agtaactccg agcattcact ctacaaatac tcaaattaat

     760     770     780     790     800
actgctaaat ctggaactat gaaccaagt agcgacgatg aggaccttat

     810     820     830     840     850
gatttccggc tcaggatctg gagaacaggg gcccacaca actcttaatg

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

[illegible]

```

ccttaatcaa tagacatgct attgactttc tactcacaag atggggagga
      1810      1820      1830      1840      1850
acatgcaaag tgcttggacc tgattgttgc atcgggatag aagacttgtc
      1860      1870      1880      1890      1900
caaaaatatt tcagagcaaa ttgaccaaat taaaaaggac gaacaaaaag
      1910      1920      1930      1940      1950
aggggactgg ttggggctctg ggtggtaa atc ggtggacatc cgactggggt
      1960      1970      1980      1990      2000
gttcttacta acttgggcat tttgctacta ttatccatag ctgtcttgat
      2010      2020      2030      2040      2050
tgctctatcc tgtatttgtc gtatctttac taaatatatc ggataa

```

SEQ ID NO:8

RVNGP1/MUSGP2 amino acid sequence

```

MetLysThrIleTyrPheLeuIleSerLeuIleLeuIleGlnSerIleLys
      ThrLeuProValLeuGluIleAlaSerAsnSerGlnProGlnAspVal
AspSerValCysSerGlyThrLeuGlnLysThrGluAspValHisLeuMet
GlyPheThrLeuSerGlyGlnLysValAlaAspSerProLeuGluAlaSer
      LysArgTrpAlaPheArgThrGlyValProProLysAsnValGluTyr
ThrGluGlyGluGluAlaLysThrCysTyrAsnIleSerValThrAspPro
SerGlyLysSerLeuLeuAspProProSerAsnIleArgAspTyrPro
      LysCysLysThrValHisHisIleGlnGlyGlnAsnProHisAlaGln
GlyIleAlaLeuHisLeuTrpGlyAlaPhePheLeuTyrAspArgValAla
SerThrThrMetTyrArgGlyLysValPheThrGluGlyAsnIleAlaAla
      MetIleValAsnLysThrValHisArgMetIlePheSerArgGlnGly
GlnGlyTyrArgHisMetAsnLeuThrSerThrAsnLysTyrTrpThrSer
SerAsnGluThrGlnArgAsnAspThrGlyCysPheGlyIleLeuGlnGlu
      TyrAsnSerThrAsnAsnGlnThrCysProProSerLeuLysProPro
SerLeuProThrValThrProSerIleHisSerThrAsnThrGlnIleAsn
ThrAlaLysSerGlyThrMetAsnProSerSerAspAspGluAspLeuMet
      IleSerGlySerGlySerGlyGluGlnGlyProHisThrThrLeuAsn
ValValThrGluGlnLysGlnSerSerThrIleLeuSerThrProSerLeu
HisProSerThrSerGlnHisGluGlnAsnSerThrAsnProSerArgHis
      AlaValThrGluHisAsnGlyThrAspProThrThrGlnProAlaThr

```

LeuLeuAsnAsnThrAsnThrThrProThrTyrAsnThrLeuLysTyrAsn
 LeuSerThrProSerProProThrArgAsnIleThrAsnAsnAspThrGln
 ArgGluLeuAlaGluSerGluGlnThrAsnAlaGlnLeuAsnThrThr
 LeuAspProThrGluAsnProThrThrGlyGlnAspThrAsnSerThrThr
 AsnIleIleMetThrThrSerAspIleThrSerLysHisProThrAsnSer
 SerProAspSerSerProThrThrArgProProIleTyrPheArgLys
 LysArgSerIleLeuTrpArgGluGlyAspMetPheProPheLeuAspGly
 LeuIleAsnAlaProIleAspPheAspProValProAsnThrLysThrIle
 PheAspGluSerSerSerSerGlyAlaSerAlaGluGluAspGlnHis
 AlaSerProAsnIleSerLeuThrLeuSerTyrPheProAsnIleAsnGlu
 AsnThrAlaTyrSerGlyGluAsnGluAsnAspCysAspAlaGluLeuArg
 IleTrpSerValGlnGluAspAspLeuAlaAlaGlyLeuSerTrpIle
 ProPhePheGlyProGlyIleGluGlyLeuTyrThrAlaValLeuIleLys
 AsnGlnAsnAsnLeuValCysArgLeuArgArgLeuAlaAsnGlnThrAla
 LysSerLeuGluLeuLeuLeuArgValThrThrGluGluArgThrPhe
 SerLeuIleAsnArgHisAlaIleAspPheLeuLeuThrArgTrpGlyGly
 ThrCysLysValLeuGlyProAspCysCysIleGlyIleGluAspLeuSer
 LysAsnIleSerGluGlnIleAspGlnIleLysLysAspGluGlnLys
 GluGlyThrGlyTrpGlyLeuGlyGlyLysTrpTrpThrSerAspTrpGly
 ValLeuThrAsnLeuGlyIleLeuLeuLeuLeuSerIleAlaValLeuIle
 AlaLeuSerCysIleCysArgIlePheThrLysTyrIleGly*

SEQ ID NO:9

MBGGP1/MBGGP2 nucleotide sequence

10	20	30	40	50
aagcttaaca tgaagaccac atgtttcctt atcagtctta tcttaattca				
60	70	80	90	100
agggacaaaa aatctcccca ttttagagat agctagtaat aatcaacccc				
110	120	130	140	150
aaaatgtgga ttcggtatgc tccggaactc tccagaagac agaagacgtc				
160	170	180	190	200
catctgatgg gattcacact gagtgggcaa aaagttgctg attccccctt				
210	220	230	240	250


```

ggaggcatcc aagcgatggg ctttcaggac aggtgtacct cccaagaatg
      260      270      280      290      300
ttgagtacac agaggggggag gaagccaaaa catgctacaa tataagtgtgta
      310      320      330      340      350
acgggatccct ctggaaaatc cttgctgtta gatcctccta ccaacatccg
      360      370      380      390      400
tgactatccg aaatgcaaaa ctatccatca tattcaagggt caaaaccctc
      410      420      430      440      450
atgcacaggg gatcgccctt catttatggg gagcattttt tctgtatgat
      460      470      480      490      500
cgcattatgt accgaggcaa agtcttcact gaagggaaca tagcagctat
      510      520      530      540      550
gattgtcaat aagacagtgc acaaaatgat tttctcgcgg caaggacaag
      560      570      580      590      600
ggtagcgtca tatgaatctg acttctacta ataaatattg gacaagtagt
      610      620      630      640      650
aacggaacgc aaacgaatga cactggatgt ttcggcgctc ttcaagaata
      660      670      680      690      700
caattctaca aagaaccaa catgtgctcc gtccaaaata cctccaccac
      710      720      730      740      750
tgcccacagc ccgtccggag atcaaactca caagcacccc aactgatgcc
      760      770      780      790      800
accaaactca ataccacgga cccaagcagt gatgatgagg acctcgcaac
      810      820      830      840      850
atccgggtca ggggtccggag aacgagaacc ccacacaact tctgatgcgg
      860      870      880      890      900
tcaccaagca agggctttca tcaacaatgc caccactcc ctcaccacaa
      910      920      930      940      950
ccaagcacgc cacagcaagg aggaaacaac acaaaccatt cccaagatgc
      960      970      980      990      1000
tgtgactgaa ctagacaaaa ataacacaac tgcacaaccg tccatgcccc
      1010      1020      1030      1040      1050
ctcataaacac taccacaatc tctactaaca acacctcaa acacaacttc
      1060      1070      1080      1090      1100
agcactctct ctgcaccatt aaaaaacacc accaatgaca acacacagag
      1110      1120      1130      1140      1150

```

cacaatcact	gaaaaatgagc	aaaccagtgcc	ccctctcgata	acaaccctgc
1160	1170	1180	1190	1200
ctccaacgggg	aaatcccacc	acagcaaaga	gcaccagcag	caaaaaaggc
1210	1220	1230	1240	1250
cccgccacaa	cggcaccaaa	cacgacaaat	gagcatttca	ccagtcctcc
1260	1270	1280	1290	1300
ccccaccccc	agctcgcactg	cacaacatct	tgtatatattc	agaagaaaagc
1310	1320	1330	1340	1350
gatcgatcct	ctggaggggaa	ggcgcacatgt	tcctttttct	ggatggggtta
1360	1370	1380	1390	1400
ataaatgctc	caattgatatt	tgacccagtt	ccaaatacaa	aaacaatctt
1410	1420	1430	1440	1450
tgatgaatcc	tctagtctctg	gtgcctcggc	tgaggaagat	caacatgcct
1460	1470	1480	1490	1500
cccccaatat	tagtttaact	ttatcttatt	ttcctaatat	aaatgagaac
1510	1520	1530	1540	1550
actgcctact	ctggagaaaa	tgagaatgat	tgtgatgcag	agttaagaat
1560	1570	1580	1590	1600
ttggagcgtt	caggaggatg	acctggccgc	agggctcagt	tggataccgt
1610	1620	1630	1640	1650
tttttggccc	tggaattgaa	ggactttaca	ctgctgtttt	aattaaaaat
1660	1670	1680	1690	1700
caaaacaatt	tggtctgcag	gttgaggcgt	ctagccaatc	aaactgccaa
1710	1720	1730	1740	1750
atccttgtaa	ctcttattga	gagtcacaac	tgaggaaaga	acattctcct
1760	1770	1780	1790	1800
taatcaatag	acatgctatt	gactttctac	tcacaagatg	gggaggaaca
1810	1820	1830	1840	1850
tgcaaagtgc	ttggacctga	ttgttgcatc	gggatagaag	acttgtccaa
1860	1870	1880	1890	1900
aaatattttca	gagcaaattg	accaaattaa	aaaggacgaa	caaaaagagg
1910	1920	1930	1940	1950
ggactgggtt	gggtctgggt	ggtaaattgg	ggacatccga	ctgggggtgt
1960	1970	1980	1990	2000
cttactaact	tgggcatttt	gctactatta	tccatagctg	tcttgattgc
2010	2020	2030	2040	2050

tctatcctgt atttgcgta tctttactaa atatatcgga taacggaatt

c

SEQ ID NO:10

MBGGP1/MBGGP2 amino acid sequence

MetLysThrThrCysPheLeuIleSerLeuIleLeuIleGlnGlyThr
LysAsnLeuProIleLeuGluIleAlaSerAsnAsnGlnProGln
AsnValAspSerValCysSerGlyThrLeuGlnLysThrGluAspVal
HisLeuMetGlyPheThrLeuSerGlyGlnLysValAlaAspSerProLeu
GluAlaSerLysArgTrpAlaPheArgThrGlyValProProLysAsn
ValGluTyrThrGluGlyGluGluAlaLysThrCysTyrAsnIleSerVal
ThrAspProSerGlyLysSerLeuLeuLeuAspProProThrAsnIleArg
AspTyrProLysCysLysThrIleHisHisIleGlnGlyGlnAsnPro
HisAlaGlnGlyIleAlaLeuHisLeuTrpGlyAlaPhePheLeuTyrAsp
ArgIleAlaSerThrThrMetTyrArgGlyLysValPheThrGluGlyAsn
IleAlaAlaMetIleValAsnLysThrValHisLysMetIlePheSer
ArgGlnGlyGlnGlyTyrArgHisMetAsnLeuThrSerThrAsnLysTyr
TrpThrSerSerAsnGlyThrGlnThrAsnAspThrGlyCysPheGlyAla
LeuGlnGluTyrAsnSerThrLysAsnGlnThrCysAlaProSerLys
IleProProProLeuProThrAlaArgProGluIleLysLeuThrSerThr
ProThrAspAlaThrLysLeuAsnThrThrAspProSerSerAspAspGlu
AspLeuAlaThrSerGlySerGlySerGlyGluArgGluProHisThr
ThrSerAspAlaValThrLysGlnGlyLeuSerSerThrMetProProThr
ProSerProGlnProSerThrProGlnGlnGlyGlyAsnAsnThrAsnHis
SerGlnAspAlaValThrGluLeuAspLysAsnAsnThrThrAlaGln
ProSerMetProProHisAsnThrThrThrIleSerThrAsnAsnThrSer
LysHisAsnPheSerThrLeuSerAlaProLeuGlnAsnThrThrAsnAsp
AsnThrGlnSerThrIleThrGluAsnGluGlnThrSerAlaProSer

Seq. ID: 10

IleThrThrLeuProProThrGlyAsnProThrThrAlaLysSerThrSer
 SerLysLysGlyProAlaThrThrAlaProAsnThrThrAsnGluHisPhe
 ThrSerProProProThrProSerSerThrAlaGlnHisLeuValTyr
 PheArgArgLysArgSerIleLeuTrpArgGluGlyAspMetPheProPhe
 LeuAspGlyLeuIleAsnAlaProIleAspPheAspProValProAsnThr
 LysThrIlePheAspGluSerSerSerSerGlyAlaSerAlaGluGlu
 AspGlnHisAlaSerProAsnIleSerLeuThrLeuSerTyrPheProAsn
 IleAsnGluAsnThrAlaTyrSerGlyGluAsnGluAsnAspCysAspAla
 GluLeuArgIleTrpSerValGlnGluAspAspLeuAlaAlaGlyLeu
 SerTrpIleProPhePheGlyProGlyIleGluGlyLeuTyrThrAlaVal
 LeuIleLysAsnGlnAsnAsnLeuValCysArgLeuArgArgLeuAlaAsn
 GlnThrAlaLysSerLeuGluLeuLeuLeuArgValThrThrGluGlu
 ArgThrPheSerLeuIleAsnArgHisAlaIleAspPheLeuLeuThrArg
 TrpGlyGlyThrCysLysValLeuGlyProAspCysCysIleGlyIleGlu
 AspLeuSerLysAsnIleSerGluGlnIleAspGlnIleLysLysAsp
 GluGlnLysGluGlyThrGlyTrpGlyLeuGlyGlyLysTrpTrpThrSer
 AspTrpGlyValLeuThrAsnLeuGlyIleLeuLeuLeuLeuSerIleAla
 ValLeuIleAlaLeuSerCysIleCysArgIlePheThrLysTyrIle
 Gly*

SEQ ID NO:11

EBOGP1/EBOGP2 nucleotide sequence

10	20	30	40	50
atggg	cgtta	cagga	atatt	gcagttac
ctt	gcgtgat	cgat	tcaagagg	ac
60	70	80	90	100
atcattc	ttt	ctttggg	taa	ttatccttt
ccaaaga	aca	ttttccat	cc	
110	120	130	140	150
cacttgg	agt	catccaca	at	agcacattac
agg	tttagtga	tg	tcgacaaa	
160	170	180	190	200
ctagttt	gtc	gtgacaaa	act	gtcatccaca
aatcaattga	gatcag	ttgg		

210	220	230	240	250
actgaatctc	gaaggggaatg	gagtggcaac	tgacgtgcc	tctgcaacta
260	270	280	290	300
aaagatgggg	cttcagggtcc	ggtgtcccac	caaaggtgg	caattatgaa
310	320	330	340	350
gctggtgaat	gggctgaaaa	ctgctacaat	cttgaaatca	aaaaaacctga
360	370	380	390	400
cgggagtgag	tgtctaccag	cagcgccaga	cgggattcgg	ggcttccccc
410	420	430	440	450
ggtgccggta	tgtgcacaaa	gtatcaggaa	cgggaccgtg	tgccggagac
460	470	480	490	500
tttgctttcc	ataaaagagg	tgcttttctt	ctgtatgatc	gacttgcttc
510	520	530	540	550
cacagttatc	taccgaggaa	cgactttcgc	tgaaggtgtc	gttgcatttc
560	570	580	590	600
tgatactgcc	ccaagctaag	aaggacttct	tcagctcaca	ccccttgaga
610	620	630	640	650
gagccgggtc	atgcaacgga	ggaccogtct	agtggctact	attctaccac
660	670	680	690	700
aattagatat	caggctaccg	gttttggaac	caatgagaca	gagtacttgt
710	720	730	740	750
tcgaggttga	caatttgacc	tacgtccaac	ttgaatcaag	attcacacca
760	770	780	790	800
cagttttctgc	tccagctgaa	tgagacaata	tatacaagt	ggaaaaggag
810	820	830	840	850
caataccacg	ggaaaactaa	tttggaaggt	caaccccgaa	attgatacaa
860	870	880	890	900
caatcggggg	gtgggccttc	tgggaaacta	aaaaaacct	cactagaaaa
910	920	930	940	950
attcgcagt	aagagttgtc	tttcacagtt	gtatcaaacg	gagccaaaaa
960	970	980	990	1000
catcagtgg	cagagtccgg	cgcgaacttc	ttocgaccca	gggaccaaca
1010	1020	1030	1040	1050
caacaactga	agaccacaaa	atcatggcct	cagaaaattc	ctctgcaatg
1060	1070	1080	1090	1100
gttcaagtgc	acagtcaagg	aagggaagct	gcagtgtcgc	atctaacaaa

1110	1120	1130	1140	1150
ccttgccaca	atctccacga	gtccccaatc	cctcacaacc	aaaccaggtc
1160	1170	1180	1190	1200
cggacaacag	cacccataat	acacccgtgt	ataaacttga	catctctgag
1210	1220	1230	1240	1250
gcaactcaag	ttgaacaaca	tcaccgcaga	acagacaacg	acagcacagc
1260	1270	1280	1290	1300
ctccgacact	ccctctgccca	cgaccgcagc	cggacccccca	aaagcagaga
1310	1320	1330	1340	1350
acaccaaacac	gagcaagagc	actgacttcc	tggacccccgc	caccacaaca
1360	1370	1380	1390	1400
agtccccaaa	accacagcga	gaccgctggc	aacaacaaca	ctcatcacca
1410	1420	1430	1440	1450
agataccgga	gaagagagtg	ccagcagcgg	gaagctaggc	ttaattacca
1460	1470	1480	1490	1500
atactattgc	tggagtcgca	ggactgatca	caggcggggag	aagaactcga
1510	1520	1530	1540	1550
cgatcggcaa	ttgtcaatgc	tcaacccaaa	tgcaacccta	atttacatta
1560	1570	1580	1590	1600
ctggactact	caggatgaag	gtgctgcaat	cggactggcc	tggataccat
1610	1620	1630	1640	1650
atttcggggcc	agcagccgag	ggaattttaca	tagaggggct	aatgcacaat
1660	1670	1680	1690	1700
caagatgggtt	taatctgtgg	gttgagacag	ctggccaacg	agacgactca
1710	1720	1730	1740	1750
agctcttcaa	ctgttctga	gagccacaac	tgagctacgc	accttttcaa
1760	1770	1780	1790	1800
tcctcaaccg	taaggcaatt	gattttcttgc	tcgagcgatg	gggcggcaca
1810	1820	1830	1840	1850
tgccacattc	tgggaccgga	ctgctgtatc	gaaccacatg	attggaccaa
1860	1870	1880	1890	1900
gaacataaca	gacaaaattg	atcagattat	tcattgatttt	gttgataaaaa
1910	1920	1930	1940	1950
cccttccgga	ccagggggac	aatgacaatt	ggtggacagg	atggagacaa
1960	1970	1980	1990	2000
tggataccgg	caggtattgg	agttacaggc	gttataattg	cagttatcgc

2010 2020 2030
tttattctgt atatgcaaatt ttgtctttta gttgaattc

SEQ ID NO:12

EBOGP1/EBOGP2 amino acid sequence

MetGlyValThrGlyIleLeuGlnLeuProArgAspArgPheLys
ArgThrSerPhePheLeuTrpValIleIleLeuPheGlnArgThrPhe
SerIleProLeuGlyValIleHisAsnSerThrLeuGlnValSerAsp
ValAspLysLeuValCysArgAspLysLeuSerSerThrAsnGlnLeuArg
SerValGlyLeuAsnLeuGluGlyAsnGlyValAlaThrAspValProSer
AlaThrLysArgTrpGlyPheArgSerGlyValProProLysValVal
AsnTyrGluAlaGlyGluTrpAlaGluAsnCysTyrAsnLeuGluIleLys
LysProAspGlySerGluCysLeuProAlaAlaProAspGlyIleArgGly
PheProArgCysArgTyrValHisLysValSerGlyThrGlyProCys
AlaGlyAspPheAlaPheHisLysGluGlyAlaPhePheLeuTyrAspArg
LeuAlaSerThrValIleTyrArgGlyThrThrPheAlaGluGlyValVal
AlaPheLeuIleLeuProGlnAlaLysLysAspPhePheSerSerHis
ProLeuArgGluProValAsnAlaThrGluAspProSerSerGlyTyrTyr
SerThrThrIleArgTyrGlnAlaThrGlyPheGlyThrAsnGluThrGlu
TyrLeuPheGluValAspAsnLeuThrTyrValGlnLeuGluSerArg
PheThrProGlnPheLeuLeuGlnLeuAsnGluThrIleTyrThrSerGly
LysArgSerAsnThrThrGlyLysLeuIleTrpLysValAsnProGluIle
AspThrThrIleGlyGluTrpAlaPheTrpGluThrLysLysAsnLeu
ThrArgLysIleArgSerGluGluLeuSerPheThrValValSerAsnGly
AlaLysAsnIleSerGlyGlnSerProAlaArgThrSerSerAspProGly
ThrAsnThrThrThrGluAspHisLysIleMetAlaSerGluAsnSer
SerAlaMetValGlnValHisSerGlnGlyArgGluAlaAlaValSerHis
LeuThrThrLeuAlaThrIleSerThrSerProGlnSerLeuThrThrLys
ProGlyProAspAsnSerThrHisAsnThrProValTyrLysLeuAsp

“EBOGP1” amino acid sequence

IleSerGluAlaThrGlnValGluGlnHisHisArgArgThrAspAsnAsp
 SerThrAlaSerAspThrProSerAlaThrThrAlaAlaGlyProProLys
 AlaGluAsnThrAsnThrSerLysSerThrAspPheLeuAspProAla
 ThrThrThrSerProGlnAsnHisSerGluThrAlaGlyAsnAsnAsnThr
 HisHisGlnAspThrGlyGluGluSerAlaSerSerGlyLysLeuGlyLeu>
 IleThrAsnThrIleAlaGlyValAlaGlyLeuIleThrGlyGlyArg
 ArgThrArgArgSerAlaIleValAsnAlaGlnProLysCysAsnProAsn
 LeuHisTyrTrpThrThrGlnAspGluGlyAlaAlaIleGlyLeuAlaTrp
 IleProTyrPheGlyProAlaAlaGluGlyIleTyrIleGluGlyLeu
 MethisAsnGlnAspGlyLeuIleCysGlyLeuArgGlnLeuAlaAsnGlu
 ThrThrGlnAlaLeuGlnLeuPheLeuArgAlaThrThrGluLeuArgThr
 PheSerIleLeuAsnArgLysAlaIleAspPheLeuLeuGlnArgTrp
 GlyGlyThrCysHisIleLeuGlyProAspCysCysIleGluProHisAsp
 TrpThrLysAsnIleThrAspLysIleAspGlnIleIleHisAspPheVal
 AspLysThrLeuProAspGlnGlyAspAsnAspAsnTrpTrpThrGly
 TrpArgGlnTrpIleProAlaGlyIleGlyValThrGlyValIleIleAla
 ValIleAlaLeuPheCysIleCysLysPheValPhe*

SEQ ID NO:13

RVNGP1/RVNGP2 nucleotide sequence

10	20	30	40	50
atgaagacca tatattttct gattagtctc attttaatcc aaagtataaa				
60	70	80	90	100
aactctccct gtttttagaaa ttgctagtaa cagccaacct caagatgtag				
110	120	130	140	150
attcagtgtg ctccggaacc ctccaaaaga cagaagatgt tcattctgatg				
160	170	180	190	200
ggatttacac tgagtgggca aaaagttgct gattcccctt tggaagcatc				
210	220	230	240	250
taaacgatgg gctttcagga caggtgttcc tccaagaac gttgagtata				

260	270	280	290	300
cggaaggaga	agaagccaaa	acatgttaca	atataagtgt	aacagaccct
310	320	330	340	350
tctggaaaat	ccttgctgct	ggatcctccc	agtaatatcc	gcgattaccc
360	370	380	390	400
taaatgtaaa	actgttcac	atattcaagg	tcaaaaccct	catgcacagg
410	420	430	440	450
ggattgccct	ccatttgtgg	ggggcatttt	tcttgtatga	tcgcgttgcc
460	470	480	490	500
tctacaacaa	tgtaccgagg	caaggtcttc	actgaaggaa	atatagcagc
510	520	530	540	550
tatgattggt	aataagacag	ttcacagaat	gattttttct	aggcaaggac
560	570	580	590	600
aaggttatcg	tcacatgaac	ttgacctcca	ccaataaata	ttggacaagc
610	620	630	640	650
agcaatgaaa	cgcagagaaa	tgatacggga	tgtttttgga	tcctccaaga
660	670	680	690	700
atacaactcc	acaaacaatc	aaacatgccc	tccatctctt	aaacctccat
710	720	730	740	750
ccctgcccac	agtaactccg	agcattcact	ctacaaatac	tcaaattaat
760	770	780	790	800
actgctaaat	ctggaactat	gaaccceaagt	agcgacgatg	aggaccttat
810	820	830	840	850
gatttccggc	tcaggatctg	gagaacaggg	gccccacaca	actcttaatg
860	870	880	890	900
tagtcaactga	acagaaacaa	tcgtcaacaa	tattgtccac	tccttcacta
910	920	930	940	950
catccaagca	cctcacaaca	tgagcaaaac	agtacgaatc	cttcccgaca
960	970	980	990	1000
tgctgtaact	gagcacaatg	gaaccgaccc	aacaacacaa	ccagcaacgc
1010	1020	1030	1040	1050
tcctcaacaa	tactaataca	actcccacct	ataacactct	caagtacaac
1060	1070	1080	1090	1100
ctcagtactc	cttcccctcc	aaccgcgaac	atcaccaata	atgatacaca
1110	1120	1130	1140	1150
acgtgaacta	gcagaaagcg	aacaaaccaa	tgctcagttg	aacacaactc

1160 1170 1180 1190 1200
 tagatccaac agaaaatccc accacaggac aagacaccaa cagcacaacc
 1210 1220 1230 1240 1250
 aacatcatca tgacgacatc agatataaca agcaaacc ccacaaattc
 1260 1270 1280 1290 1300
 ttctccggat tctagtccga caaccgccc tcctatatac tttagaaaga
 1310 1320 1330 1340 1350
 aacgatcgat tttctggaaa gaagggtgata tattcccggt tttagatggg
 1360 1370 1380 1390 1400
 ttaataaata ctgaaattga ttttgatcca atcccaaaca cagaaacaat
 1410 1420 1430 1440 1450
 ctttgatgaa tctcccagct ttaatacttc aactaatgag gaacaacaca
 1460 1470 1480 1490 1500
 ctccccgaa tatcagttta actttctctt attttctga taaaaatgga
 1510 1520 1530 1540 1550
 gatactgcct actctgggga aaacgagaat gattgtgatg cagagttgag
 1560 1570 1580 1590 1600
 gatttggagt gtgcaggagg acgatttggc ggcagggctt agctggatac
 1610 1620 1630 1640 1650
 catttttttg ccctggaatc gaaggactct atactgccgg tttaatcaaa
 1660 1670 1680 1690 1700
 aatcagaaca atttagtttg taggttgagg cgcttagcta atcaaactgc
 1710 1720 1730 1740 1750
 taaatccttg gagctcttgt taagggtcac aaccgaggaa aggacatttt
 1760 1770 1780 1790 1800
 ccttaatcaa taggcatgca attgactttt tgcttacgag gtggggcgga
 1810 1820 1830 1840 1850
 acatgcaagg tgctaggacc tgattgttgc ataggaatag aagatctatc
 1860 1870 1880 1890 1900
 taaaaatata tcagaacaaa tcgacaaaat cagaaaggat gaacaaaagg
 1910 1920 1930 1940 1950
 aggaaactgg ctgggggtcta ggtggcaa at ggtggacatc tgactggggt
 1960 1970 1980 1990 2000
 gttctcacca atttgggcat cctgctacta ttatctatag ctgttctgat
 2010 2020 2030 2040
 tgctctgtcc tgtatctgtc gtatcttcac taaatacatt ggatga

RVNGP1/RVNGP2 amino acid sequence

MetLysThrIleTyrPheLeuIleSerLeuIleLeuIleGlnSerIleLys
ThrLeuProValLeuGluIleAlaSerAsnSerGlnProGlnAspVal
AspSerValCysSerGlyThrLeuGlnLysThrGluAspValHisLeuMet
GlyPheThrLeuSerGlyGlnLysValAlaAspSerProLeuGluAlaSer
LysArgTrpAlaPheArgThrGlyValProProLysAsnValGluTyr
ThrGluGlyGluGluAlaLysThrCysTyrAsnIleSerValThrAspPro
SerGlyLysSerLeuLeuLeuAspProProSerAsnIleArgAspTyrPro
LysCysLysThrValHisHisIleGlnGlyGlnAsnProHisAlaGln
GlyIleAlaLeuHisLeuTrpGlyAlaPhePheLeuTyrAspArgValAla
SerThrThrMetTyrArgGlyLysValPheThrGluGlyAsnIleAlaAla
MetIleValAsnLysThrValHisArgMetIlePheSerArgGlnGly
GlnGlyTyrArgHisMetAsnLeuThrSerThrAsnLysTyrTrpThrSer
SerAsnGluThrGlnArgAsnAspThrGlyCysPheGlyIleLeuGlnGlu
TyrAsnSerThrAsnAsnGlnThrCysProProSerLeuLysProPro
SerLeuProThrValThrProSerIleHisSerThrAsnThrGlnIleAsn
ThrAlaLysSerGlyThrMetAsnProSerSerAspAspGluAspLeuMet
IleSerGlySerGlySerGlyGluGlnGlyProHisThrThrLeuAsn
ValValThrGluGlnLysGlnSerSerThrIleLeuSerThrProSerLeu
HisProSerThrSerGlnHisGluGlnAsnSerThrAsnProSerArgHis
AlaValThrGluHisAsnGlyThrAspProThrThrGlnProAlaThr
LeuLeuAsnAsnThrAsnThrThrProThrTyrAsnThrLeuLysTyrAsn
LeuSerThrProSerProProThrArgAsnIleThrAsnAsnAspThrGln
ArgGluLeuAlaGluSerGluGlnThrAsnAlaGlnLeuAsnThrThr
LeuAspProThrGluAsnProThrThrGlyGlnAspThrAsnSerThrThr
AsnIleIleMetThrThrSerAspIleThrSerLysHisProThrAsnSer
SerProAspSerSerProThrThrArgProProIleTyrPheArgLys
LysArgSerIlePheTrpLysGluGlyAspIlePheProPheLeuAspGly
LeuIleAsnThrGluIleAspPheAspProIleProAsnThrGluThrIle
PheAspGluSerProSerPheAsnThrSerThrAsnGluGluGlnHis
ThrProProAsnIleSerLeuThrPheSerTyrPheProAspLysAsnGly

AspThrAlaTyrSerGlyGluAsnGluAsnAspCysAspAlaGluLeuArg
 IleTrpSerValGlnGluAspAspLeuAlaAlaGlyLeuSerTrpIle
 ProPhePheGlyProGlyIleGluGlyLeuTyrThrAlaGlyLeuIleLys
 AsnGlnAsnAsnLeuValCysArgLeuArgArgLeuAlaAsnGlnThrAla
 LysSerLeuGluLeuLeuLeuArgValThrThrGluGluArgThrPhe
 SerLeuIleAsnArgHisAlaIleAspPheLeuLeuThrArgTrpGlyGly
 ThrCysLysValLeuGlyProAspCysCysIleGlyIleGluAspLeuSer
 LysAsnIleSerGluGlnIleAspLysIleArgLysAspGluGlnLys
 GluGluThrGlyTrpGlyLeuGlyGlyLysTrpTrpThrSerAspTrpGly
 ValLeuThrAsnLeuGlyIleLeuLeuLeuLeuSerIleAlaValLeuIle
 AlaLeuSerCysIleCysArgIlePheThrLysTyrIleGly*

SEQ ID NO:15

GP1 forward (containing HindIII) EBOV (5'-CAA GCT TCA ATG GGC GTT ACA-3'),

SEQ ID NO:16

MBGV (5'-AAG CTT AAC ATG AAG ACC ACA T-3');

SEQ ID NO:17

GP2 forward (containing PvuI) EBOV (5'-GAC GAT CGG CAA TTG TCA ATG-3'),

SEQ ID NO:18

MBGV (5'-AGC GAT CGA TCC TCT GGA G -3');

SEQ ID NO:19

GP1 reverse (containing PvuI) EBOV (5'-GCC GAT CGT CGA GTT CTT CT-3'),

SEQ ID NO:20

MBGV (5'-GAT CGA TCG CTT TCT TCT G-3');

SEQ ID NO:21

GP2 reverse (containing EcoRI) EBOV (5'-TGA ATT CAA CTA AAA GAC AAA TTT G-3'),

SEQ ID NO:22

MBGV (5'-CGA ATT CCG TTA TCC GAT ATA T-3')

SEQ ID NO:23

GP1 forward (containing HindIII) MBGV MUS (5'-AAG CTT AAC ATG AAG ACC ACA T-3')

SEQ ID NO:24

MBGV RVN (5-AAG CTT CGA CAT GAA GAC CAT AT-3');

2017-11-09 14:00:00

SEQ ID NO:25

GP2 forward (containing PvuI) MBGV MUS (5'-AGC GAT CGA TCC
TCT GGA G -3'),

SEQ ID NO:26

MBGV RVN (5'-AAC GAT CGA TTT TCT GGA A-3');

SEQ ID NO:27

GP1 reverse (containing PvuI) MBGV MUS (5'-GAT CGA TCG CTT
TCT TCT G-3'),

SEQ ID NO:28

MBGV RVN (5'-AAA TCG ATC GTT TCT TTC TAA AG-3');

SEQ ID NO:29

GP2 reverse (containing EcoRI) MBGV MUS (5'-CGA ATT CCG TTA
TCC GAT ATA T-3'),

SEQ ID NO:30

MBGV RVN (5'-CGA ATT CTG TCA TCC AAT GTA T-3').

2025-11-11 14:11:11